output3-continents-ab

Yidi

2020/8/19

## Load the packages

Load some packages for manipulating and modelling the data

library(maps)  
library(devtools)  
library(predictsFunctions)  
library(StatisticalModels)  
library(raster)  
library(dplyr)  
library(tidyr)  
library(lme4)  
library(car)  
library(DHARMa)  
library(MuMIn)  
library(Hmisc)  
library(ggplot2)  
library(gstat)  
library(scatterplot3d)

## Read in and process the diversity data

diversity <- readRDS("/Users/dd/Desktop/PREDICTS data/database.rds")  
diversity <- mutate(diversity,   
 Measurement = Effort\_corrected\_measurement,  
 Sampling\_effort = Rescaled\_sampling\_effort)

An optional step to merge any sites that are within the same land-use type and that have identical coordinates, start and end dates.

diversity <- MergeSites(diversity, silent = TRUE)

## Calculate the diversity metrics

sites <- diversity %>%  
   
 # add Diversity\_metric\_is\_valid column  
 mutate(Diversity\_metric\_is\_valid = TRUE) %>%  
   
 # calculate SiteMetrics   
 SiteMetrics(extra.cols = c("SSB", "SSBS", "Predominant\_land\_use", "UN\_region")) %>%  
   
 # calculate the total abundance within each study  
 group\_by(SS) %>%  
 mutate(MaxAbundance = ifelse(Diversity\_metric\_type == "Abundance",  
 max(Total\_abundance),  
 NA)) %>%  
 ungroup() %>%  
   
 # now calculate the rescaled abundance (abundance divided by the maximum within each study)  
 mutate(RescaledAbundance = ifelse(Diversity\_metric\_type == "Abundance",  
 Total\_abundance/MaxAbundance,  
 NA))

## Computing site metrics for 2906994 measurements  
## The data contain 480 sources, 666 studies and 22678 sites  
## Computing site-level values  
## Computing total abundance  
## Computing species richness  
## Assembling site-level values

sites <- sites %>%  
   
 mutate(  
   
 # collapse primary forest and non-forest together into primary vegetation as these aren't well distinguished  
 Predominant\_land\_use = recode\_factor(Predominant\_land\_use,   
 "Primary forest" = "Primary vegetation",   
 "Primary non-forest" = "Primary vegetation"),  
   
 # indeterminate secondary veg and cannot decide get NA  
 Predominant\_land\_use = na\_if(Predominant\_land\_use, "Secondary vegetation (indeterminate age)"),  
 Predominant\_land\_use = na\_if(Predominant\_land\_use, "Cannot decide"),  
   
 # set reference levels  
 Predominant\_land\_use = factor(Predominant\_land\_use),  
 Predominant\_land\_use = relevel(Predominant\_land\_use, ref = "Primary vegetation"),  
 )

## Model site-level diversity

Step 1: collinearity

source("/Users/dd/HighstatLib10.R")  
corvif(sites[ , c("Predominant\_land\_use", "UN\_region")])

##   
##   
## Variance inflation factors  
##   
## GVIF Df GVIF^(1/2Df)  
## Predominant\_land\_use 1.228252 7 1.014793  
## UN\_region 1.228252 4 1.026032

Step 2: complete cases

model\_data <- drop\_na(sites, Total\_abundance, Predominant\_land\_use, UN\_region)

Step 3: starting/maximal model

transforming RescaledAbundance

model\_data <- mutate(model\_data,   
 logAbundance = log(RescaledAbundance + 1),  
 sqrtAbundance = sqrt(RescaledAbundance))  
m1 <- lmer(sqrtAbundance ~ Predominant\_land\_use\*UN\_region +  
 (1|SS) + (1|SSB), data = model\_data)

Step 4: Choose the random effects

m2 <- lmer(sqrtAbundance ~ Predominant\_land\_use\*UN\_region +  
 (1|SS) + (1|SSB) + (1|Source\_ID), data = model\_data)

# compare the models using Akaike's Information Criterion (AIC)  
AIC(m1,m2)

## df AIC  
## m1 43 -7215.511  
## m2 44 -7233.171

Step 5: Choose the best fixed effects structure

# have a look at the significance of the terms  
Anova(m2)

## Analysis of Deviance Table (Type II Wald chisquare tests)  
##   
## Response: sqrtAbundance  
## Chisq Df Pr(>Chisq)   
## Predominant\_land\_use 155.371 7 <2e-16 \*\*\*  
## UN\_region 3.656 4 0.4546   
## Predominant\_land\_use:UN\_region 377.486 28 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# look at the model estimates of our mimumum adequate model (which in our case is also our maximal model)  
summary(m2)

## Linear mixed model fit by REML ['lmerMod']  
## Formula: sqrtAbundance ~ Predominant\_land\_use \* UN\_region + (1 | SS) +   
## (1 | SSB) + (1 | Source\_ID)  
## Data: model\_data  
##   
## REML criterion at convergence: -7321.2  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.8109 -0.5869 -0.0450 0.5526 5.0982   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## SSB (Intercept) 0.006497 0.08061   
## SS (Intercept) 0.016154 0.12710   
## Source\_ID (Intercept) 0.017090 0.13073   
## Residual 0.033377 0.18269   
## Number of obs: 17593, groups: SSB, 1937; SS, 575; Source\_ID, 424  
##   
## Fixed effects:  
## Estimate  
## (Intercept) 0.6480533  
## Predominant\_land\_useYoung secondary vegetation -0.0662870  
## Predominant\_land\_useIntermediate secondary vegetation 0.0026872  
## Predominant\_land\_useMature secondary vegetation -0.0918454  
## Predominant\_land\_usePlantation forest 0.0001211  
## Predominant\_land\_usePasture -0.1760221  
## Predominant\_land\_useCropland -0.1117158  
## Predominant\_land\_useUrban 0.0425651  
## UN\_regionAmericas 0.0092517  
## UN\_regionAsia 0.0446920  
## UN\_regionEurope 0.0728339  
## UN\_regionOceania 0.0046042  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionAmericas 0.0789435  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionAmericas 0.0151222  
## Predominant\_land\_useMature secondary vegetation:UN\_regionAmericas 0.1500827  
## Predominant\_land\_usePlantation forest:UN\_regionAmericas -0.0325084  
## Predominant\_land\_usePasture:UN\_regionAmericas 0.1772838  
## Predominant\_land\_useCropland:UN\_regionAmericas 0.1115752  
## Predominant\_land\_useUrban:UN\_regionAmericas -0.0993849  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionAsia 0.0185030  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionAsia -0.0357143  
## Predominant\_land\_useMature secondary vegetation:UN\_regionAsia 0.1048597  
## Predominant\_land\_usePlantation forest:UN\_regionAsia -0.1161642  
## Predominant\_land\_usePasture:UN\_regionAsia 0.0611000  
## Predominant\_land\_useCropland:UN\_regionAsia 0.0313407  
## Predominant\_land\_useUrban:UN\_regionAsia -0.0186969  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionEurope -0.0079921  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionEurope -0.2264547  
## Predominant\_land\_useMature secondary vegetation:UN\_regionEurope 0.0232062  
## Predominant\_land\_usePlantation forest:UN\_regionEurope -0.1095919  
## Predominant\_land\_usePasture:UN\_regionEurope 0.1073560  
## Predominant\_land\_useCropland:UN\_regionEurope -0.0443596  
## Predominant\_land\_useUrban:UN\_regionEurope -0.2341517  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionOceania -0.0044933  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionOceania -0.0317953  
## Predominant\_land\_useMature secondary vegetation:UN\_regionOceania 0.1142958  
## Predominant\_land\_usePlantation forest:UN\_regionOceania 0.0253235  
## Predominant\_land\_usePasture:UN\_regionOceania 0.0958722  
## Predominant\_land\_useCropland:UN\_regionOceania 0.1792891  
## Predominant\_land\_useUrban:UN\_regionOceania -0.1232355  
## Std. Error  
## (Intercept) 0.0251600  
## Predominant\_land\_useYoung secondary vegetation 0.0176425  
## Predominant\_land\_useIntermediate secondary vegetation 0.0177614  
## Predominant\_land\_useMature secondary vegetation 0.0274727  
## Predominant\_land\_usePlantation forest 0.0130194  
## Predominant\_land\_usePasture 0.0173630  
## Predominant\_land\_useCropland 0.0122748  
## Predominant\_land\_useUrban 0.0356997  
## UN\_regionAmericas 0.0300007  
## UN\_regionAsia 0.0352288  
## UN\_regionEurope 0.0344084  
## UN\_regionOceania 0.0402335  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionAmericas 0.0220713  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionAmericas 0.0213834  
## Predominant\_land\_useMature secondary vegetation:UN\_regionAmericas 0.0325784  
## Predominant\_land\_usePlantation forest:UN\_regionAmericas 0.0186913  
## Predominant\_land\_usePasture:UN\_regionAmericas 0.0198622  
## Predominant\_land\_useCropland:UN\_regionAmericas 0.0177157  
## Predominant\_land\_useUrban:UN\_regionAmericas 0.0410818  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionAsia 0.0273803  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionAsia 0.0286446  
## Predominant\_land\_useMature secondary vegetation:UN\_regionAsia 0.0353077  
## Predominant\_land\_usePlantation forest:UN\_regionAsia 0.0199108  
## Predominant\_land\_usePasture:UN\_regionAsia 0.0719699  
## Predominant\_land\_useCropland:UN\_regionAsia 0.0292025  
## Predominant\_land\_useUrban:UN\_regionAsia 0.0862385  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionEurope 0.0282727  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionEurope 0.0242123  
## Predominant\_land\_useMature secondary vegetation:UN\_regionEurope 0.0354516  
## Predominant\_land\_usePlantation forest:UN\_regionEurope 0.0254316  
## Predominant\_land\_usePasture:UN\_regionEurope 0.0226575  
## Predominant\_land\_useCropland:UN\_regionEurope 0.0233578  
## Predominant\_land\_useUrban:UN\_regionEurope 0.0464009  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionOceania 0.0262040  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionOceania 0.0330313  
## Predominant\_land\_useMature secondary vegetation:UN\_regionOceania 0.0769831  
## Predominant\_land\_usePlantation forest:UN\_regionOceania 0.0355844  
## Predominant\_land\_usePasture:UN\_regionOceania 0.0230459  
## Predominant\_land\_useCropland:UN\_regionOceania 0.0344463  
## Predominant\_land\_useUrban:UN\_regionOceania 0.0587526  
## t value  
## (Intercept) 25.757  
## Predominant\_land\_useYoung secondary vegetation -3.757  
## Predominant\_land\_useIntermediate secondary vegetation 0.151  
## Predominant\_land\_useMature secondary vegetation -3.343  
## Predominant\_land\_usePlantation forest 0.009  
## Predominant\_land\_usePasture -10.138  
## Predominant\_land\_useCropland -9.101  
## Predominant\_land\_useUrban 1.192  
## UN\_regionAmericas 0.308  
## UN\_regionAsia 1.269  
## UN\_regionEurope 2.117  
## UN\_regionOceania 0.114  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionAmericas 3.577  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionAmericas 0.707  
## Predominant\_land\_useMature secondary vegetation:UN\_regionAmericas 4.607  
## Predominant\_land\_usePlantation forest:UN\_regionAmericas -1.739  
## Predominant\_land\_usePasture:UN\_regionAmericas 8.926  
## Predominant\_land\_useCropland:UN\_regionAmericas 6.298  
## Predominant\_land\_useUrban:UN\_regionAmericas -2.419  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionAsia 0.676  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionAsia -1.247  
## Predominant\_land\_useMature secondary vegetation:UN\_regionAsia 2.970  
## Predominant\_land\_usePlantation forest:UN\_regionAsia -5.834  
## Predominant\_land\_usePasture:UN\_regionAsia 0.849  
## Predominant\_land\_useCropland:UN\_regionAsia 1.073  
## Predominant\_land\_useUrban:UN\_regionAsia -0.217  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionEurope -0.283  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionEurope -9.353  
## Predominant\_land\_useMature secondary vegetation:UN\_regionEurope 0.655  
## Predominant\_land\_usePlantation forest:UN\_regionEurope -4.309  
## Predominant\_land\_usePasture:UN\_regionEurope 4.738  
## Predominant\_land\_useCropland:UN\_regionEurope -1.899  
## Predominant\_land\_useUrban:UN\_regionEurope -5.046  
## Predominant\_land\_useYoung secondary vegetation:UN\_regionOceania -0.171  
## Predominant\_land\_useIntermediate secondary vegetation:UN\_regionOceania -0.963  
## Predominant\_land\_useMature secondary vegetation:UN\_regionOceania 1.485  
## Predominant\_land\_usePlantation forest:UN\_regionOceania 0.712  
## Predominant\_land\_usePasture:UN\_regionOceania 4.160  
## Predominant\_land\_useCropland:UN\_regionOceania 5.205  
## Predominant\_land\_useUrban:UN\_regionOceania -2.098

##   
## Correlation matrix not shown by default, as p = 40 > 12.  
## Use print(x, correlation=TRUE) or  
## vcov(x) if you need it

## Plot the results

source("/Users/dd/Desktop/R script/PlotErrBar\_interactions.R")  
source("/Users/dd/Desktop/R script/J-Yidi.R")

PlotErrBar\_interactions(model = m2  
 , resp = "sqrt(Abundance)"  
 , Effect1 = "Predominant\_land\_use"  
 , Effect2 = "UN\_region"  
 , off= -0.4  
 , off\_increment = 0.2  
 , legend = TRUE  
 , leg.pos = "topleft"  
 , blackwhite = FALSE  
 , ylims = c(-0.4,0.4)  
 , srttxt =30)

## Loading required package: roquefort

## Warning in library(package, lib.loc = lib.loc, character.only = TRUE,  
## logical.return = TRUE, : there is no package called 'roquefort'

